Multimedia Systems
Lecture – 19

By

Dr. Priyambada Subudhi
Assistant Professor
IIIT Sri City

Differential Pulse Code Modulation (DPCM)

- It is based on differential predictive coding.
- Audio is often stored not in simple PCM but in a form that exploits differences.
- Generally, if a time-dependent signal has some consistency over time (*temporal redundancy*), the difference signal—subtracting the current sample from the previous one—will have a more peaked histogram, with a maximum around zero.
- For a start, differences will generally be smaller numbers and hence offer the possibility of using fewer bits to store.

- Suppose our integer sample values are in the range 0 .. 255. Then differences could be as much as −255 .. 255. So we have unfortunately increased our *dynamic range* (ratio of maximum to minimum) by a factor of two.
- Let's formalize our statement of what we are doing by defining the integer signal as the set of values f_n.
- Then we *predict* $\hat{f}_n$ values as simply the previous value, and we define the error e_n as the difference between the actual and predicted signals:

$$\hat{f}_n = f_{n-1}$$
$$e_n = f_n - \hat{f}_n$$

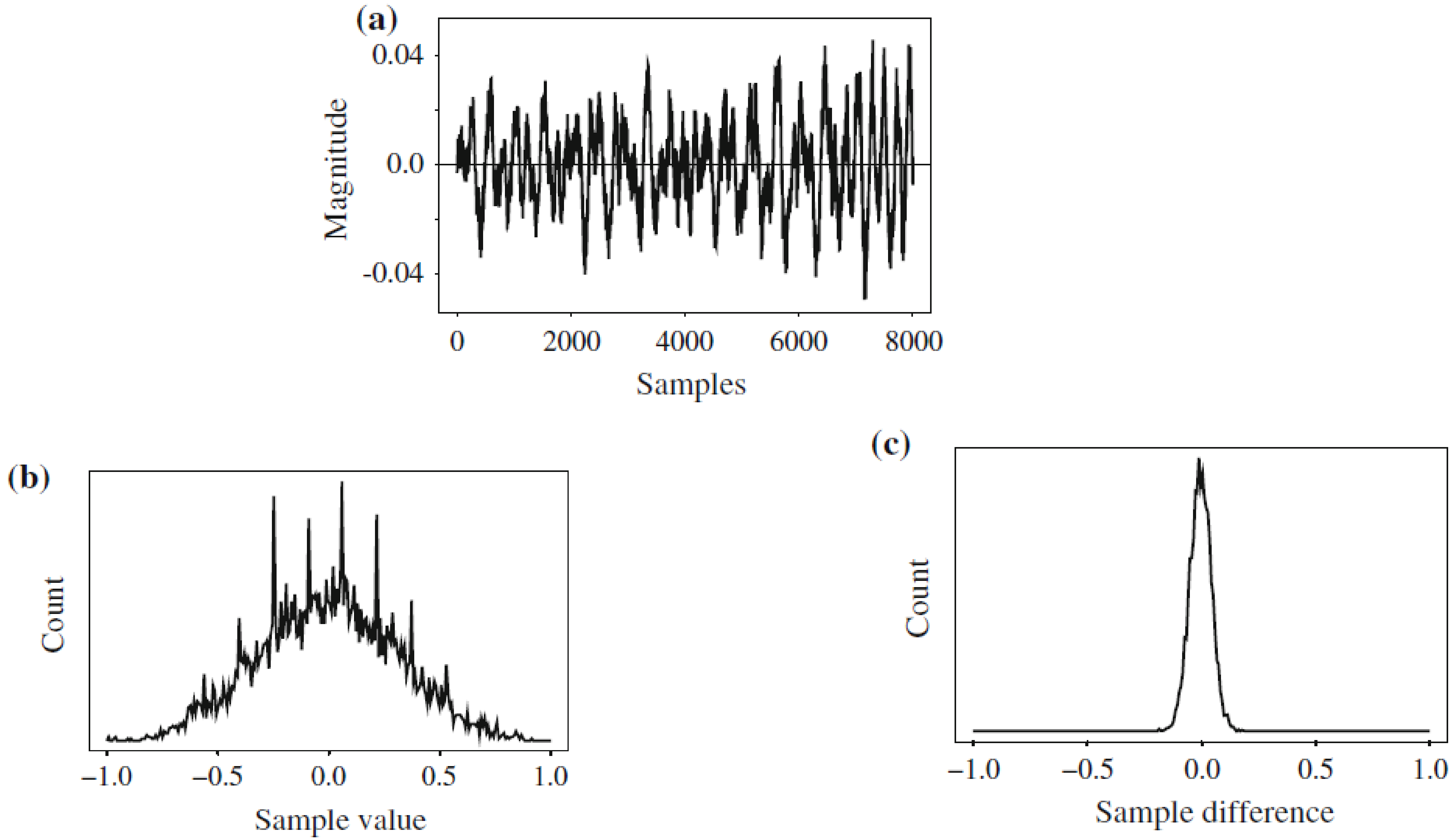
- We certainly would like our error value e_n to be as small as possible.
- Therefore, we would wish our prediction $\hat{f}_n$ to be as close as possible to the actual signal f_n.

- But for a particular sequence of signal values, some *function* of a few of the previous values, f_{n-1}, f_{n-2}, f_{n-3}, etc., may provide a better prediction of f_n.

$$\hat{f}_{\mathrm{n}} = \sum_{k=1}^{2 \text{ to } 4} a_{n-k} f_{n-k}$$

- The idea of forming differences is to make the histogram of sample values more peaked.
- A plot of 1 second of sampled speech at 8 kHz, with magnitude resolution of 8 bits per sample is considered.
- A histogram of these values is centered around zero. However, the histogram for corresponding speech signal *differences*: difference values are much more clustered around zero than are sample values themselves.
- So we can assign short codes to prevalent values like zeros here and long codewords to rarely occurring ones.

Differencing concentrates the histogram: **a** digital speech signal; **b** histogram of digital speech signal values; **c** histogram of digital speech signal differences



- Suppose samples are in the range 0 .. 255, and differences are in −255 .. 255. Then
- Define SU and SD as shifts by 32.
- Then we could in fact produce code-words for a limited set of signal differences, say only the range −15 .. 16.
- Differences (that inherently are in the range −255 .. 255) lying in the limited range can be coded as is, but if we add the extra two values for SU, SD, a value outside the range −15 .. 16 can be transmitted as a series of shifts, followed by a value that is indeed inside the range −15 .. 16.
- For example, 100 is transmitted as SU, SU, SU, 4, where (the codes for) SU and for 4 are what are sent.

- Example: suppose we devise a predictor for $\hat{f}_n$ as follows:

$$\hat{f}_n = \lfloor \frac{1}{2}(f_{n-1} + f_{n-2}) \rfloor$$
$$e_n = f_n - \hat{f}_n$$

- Then the error e_n (or a codeword for it) is what is actually transmitted. Suppose we wish to code the sequence $f1, f2, f3, f4, f5$ = 21, 22, 27, 25, 22.
- For the purposes of the predictor, we'll invent an extra signal value f_0, equal to $f1$ = 21, and first transmit this initial value, uncoded; after all, every coding scheme has the extra expense of some header information.

- Then the first error, *e*1, is zero, and subsequently

$$\hat{f}_2 = 21, \quad e_2 = 22 - 21 = 1$$

$$\hat{f}_3 = \lfloor\frac{1}{2}(f_2 + f_1)\rfloor = \lfloor\frac{1}{2}(22 + 21)\rfloor = 21$$
$$e_3 = 27 - 21 = 6$$

$$\hat{f}_4 = \lfloor\frac{1}{2}(f_3 + f_2)\rfloor = \lfloor\frac{1}{2}(27 + 22)\rfloor = 24$$
$$e_4 = 25 - 24 = 1$$

$$\hat{f}_5 = \lfloor\frac{1}{2}(f_4 + f_3)\rfloor = \lfloor\frac{1}{2}(25 + 27)\rfloor = 26$$
$$e_5 = 22 - 26 = -4$$

Schematic diagram for Predictive Coding: **a** encoder; **b** decoder